

‘Environment’ submissions in the UK’s Research Excellence Framework 2014

Article (Accepted Version)

Thorpe, A, Craig, R, Tourish, D, Hadikin, G and Batistic, S (2018) ‘Environment’ submissions in the UK’s Research Excellence Framework 2014. *British Journal of Management*, 29 (3). pp. 571-587. ISSN 1045-3172

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/68710/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

‘Environment’ Submissions in the UK’s Research Excellence Framework 2014

There has been much debate about university research assessment exercises. In the UK, a major element of the 2014 Research Excellence Framework (REF2014) has been the research ‘Environment’. Here we analyse 98 REF2014 ‘Environment’ submissions in Business and Management Studies. We explore whether there are distinctive language-related differences between submissions of high and low ranked universities, and conclude that submission writers have a strong incentive to exaggerate strengths and conceal problems. In addition, innate biases such as the ‘halo’ and ‘velcro’ effects may distract the attention of assessors from a submission’s strengths and weaknesses, since they are likely to influence their pre-existing impressions. We propose several changes to improve how ‘Environment’ is evaluated. We also argue that the research ‘Environment’ would be more likely to be enhanced if the number of outputs submitted in future were an average of two and a maximum of four per academic, rather than the maximum of six currently being considered.

Keywords: REF, Impression Management, Language, Policy Implications

Introduction

Linking research funding to evaluations of the perceived quality of research reflects embrace of a ‘New Public Management’ (NPM) mentality by higher education policy makers (Craig *et al.*, 2014). NPM assumes that the promotion of markets, managers and measurement improves performance in the public sector (Ferlie *et al.*, 1996). In the higher education sector, the NPM approach assumes that the quality of research can be quantified and measured accurately. We subject this assumption to critical inquiry by focusing on submissions regarding (research) ‘Environment’ that were made in the UK’s Research Evaluation Framework in 2014 (REF2014).

The need for such analysis arises from the growing complexity and costs imposed by successive research assessment exercises. REF2014 was estimated to cost £250 million (*Times Higher Education*, 2015). Thirty-six expert panels completed peer reviews in their respective ‘Unit of Assessment’ (UoA) of submissions regarding ‘Outputs’ (65%), ‘Impact’ (20%), and ‘Environment’ (15%).

We focus on the language used in the ‘Environment’ element of REF2014 in submissions to UoA19, Business and Management Studies [B&M]. Although the ‘Impact’ component of REF2014 has been subjected to detailed scholarly scrutiny (Manville et al. 2015; Derrick and Samuel, 2016; Kellard and Sliwa, 2016), the ‘Environment’ element has not. Here we explore how research ‘Environment’ submissions extolled virtues, minimised difficulties, inflated performance outcomes, and crafted what their authors hoped was a compelling story. Our analysis supports emerging critiques of research assessment exercises for inaccurately measuring quality, costs and effects. Those critiques have also explored how assessment exercises might be improved (Saunders et al. 2011; Linkova, 2014; Mingers and White, 2015; Geuna and Piolatto, 2016).

We were motivated to conduct the present study by three factors. First, was the disturbing conclusion of assessor bias by Taylor (2011, p.211) with respect to the Research Assessment Exercise [RAE] 2008 for ‘Research environment and esteem’ in UoAs for Accounting and Finance, Business and Management, and Economics and Econometrics.¹ We were curious to explore whether the same alleged bias persisted in the REF2014 assessment of ‘Environment.’ Our curiosity was piqued by knowledge that 12 of the 17 academic assessors for RAE 2008 were among the 24 academic assessors in B&M for REF2014.

Second, a Higher Education Funding Council for England (HEFCE) report expressed concerns that ‘...the narrative elements [of REF2014 submissions] were hard to assess, with difficulties in separating quality in research environment from quality in writing about it’ (Wilsdon *et al.*, 2015. p.129). This concern was reiterated forcefully in the consultation document on the next UK REF exercise (HEFCE, 2016, note 112).

A third motivator was the extent to which quantified metrics appeared to influence REF2014 B&M ‘Environment’ scores. We performed a multivariate regression (OLS) analysis, similar to that of Taylor (2011), with the ‘Environment’ GPA score as the dependent variable. Like Taylor, our explanatory variables included size (Full Time Equivalent [FTE] staff submitted), research income per FTE, number of postgraduate research degree [PGR] completions per FTE. We also inserted dummy variables to control for whether the HEI was a member of the Russell Group of Universities (designated as ‘Russell’), or had an assessor (‘Assessor’) on the panel.

Consistent with Taylor (2011), we found that PGR completions had no significant effect on GPA scores (see Appendix, Model 1). The model of best fit (Appendix, Model 2) revealed

no evidence of multicollinearity. This model included FTEs submitted, research income per FTE, Russell, and Assessor. Our findings mirror Taylor's, inasmuch as both Income per FTE and FTE submitted were highly significant (1% level), but diverged in terms of the membership dummies. In RAE2008, Taylor (2011) found the effects of Russell Group membership was also highly significant (1% level) and generated a GPA premium of 0.43 points. This led him to speculate that such universities may have benefited from 'a "halo effect" independent of their recent research activity' (p.214). Our research indicates that Russell Group membership was less important and the GPA premium considerably less (0.21 points) in REF2014. In contrast, while Taylor (2011, p.211) reported 'zero panel membership bias across all three research profiles [UoAs]' in RAE2008, we found having a panel member was more strongly significant, contributing to an increase of 0.375 in recorded GPA.²

Importantly, the explanatory power of our regression (62.9%) suggests the accompanying narrative played an important role in determining REF2014 environment scores. This finding prompted us to study the role of narrative in influencing research 'Environment' scores. The research question we therefore explore is:

Are REF2014 'Environment' submissions in Business and Management Studies distinctively different in language-related characteristics between high-ranked and low-ranked university submissions?

This question has important implications. *If* assessment ratings are associated positively with narrative, this should invite re-thinking of whether to include 'Environment' submissions (in their current form) in future research assessment exercises.

We begin by noting the strong parallels between university impression management (exercised through 'Environment' submissions) and methods of corporate reputation building through language use. We then explain our research method, before discussing results. We conclude that strong consideration should be given to re-thinking how 'Environment' is assessed in the future and offer some policy recommendations to help promote debate.

Impression management and language choice

The relationship between language choice and corporate reputation has been explored extensively in marketing and business management (e.g., Amernic and Craig, 2007; Geppert and Lawrence, 2008). Impression management theory explains that text can be manipulated deliberately by techniques that include making the reading difficult, cultivating an optimistic

tone, and using a complex physical layout. Manipulation can be for positive or negative rhetorical purpose (Cho *et al.*, 2010). However, the relationship between language choice and reputation has been ignored in analyses of REF2014 and RAE submissions. Debate about the accountability of public institutions has elicited concerns about image management by HEIs (Stein, 1990). In the UK, pressures to manage institutional image have been intensified by the proliferation of performance league tables. The management of institutional image has become critical to ‘the competitiveness of HEI’ (Duarte *et al.*, 2010, p.21). Similar trends are evident elsewhere (Slaughter and Rhoades, 2009; Côté and Allahar, 2011). Universities aim to present a ‘polished, unified media image’ (Brass and Rowe (2009, p.53; see also Wernick, 2006). In doing so, they are alleged to exaggerate employment prospects, academic quality, and the pleasures of university social life (Duarte *et al.*, 2010; Matherly, 2012). In the UK, they project desirable images by promoting their membership of university ‘mission groups’. The Russell Group, for example, has positioned itself in the public psyche as encompassing *the* research elite in the UK by claiming to represent 24 ‘research intensive world-class universities’ (<http://www.russellgroup.ac.uk/about/our-universities/>). This gives its members considerable reputational cache. When results of REF2014 were announced, many university websites presented their results in the most positive light possible, consistent with the view that business schools are forced ‘to play a game of illusion, to choose to misrepresent themselves’ (Roper and Davies, 2007, p.76). More recently, for example, the University of Reading was ‘forced to withdraw its assertion that it was in the top 1% of institutions globally’ following complaints made to the Advertising Standards Authority, in April 2017 (The Guardian, June 8 2017) ³.

Wittingly or otherwise, universities may be seeking to use impression management techniques to build up a ‘halo effect’. This has been described as ‘one of the oldest and most widely known of psychological phenomena’ (Nisbett and Wilson, 1977, p.250). Its origins are attributed to the work of Thorndike (1920), who noticed a tendency for soldiers who were rated positively on one or a small number of characteristics to be rated favourably overall. We consider whether the halo effect is pertinent in the context of assessments of REF2014 ‘Environment.’ Thorndike (1920) argued that the halo effect represented a fundamental inability to resist the affective influence [that is, of feelings or emotions] in making global evaluations of specific attributes (Nisbett and Wilson, 1977).

Taylor (2011, p.14) suggested that RAE2008 panel members may have suffered from implicit bias when making judgements on the research environment of Russell Group members

due to the ‘halo effect’. Given that over two-thirds (12 of the 17) of the UoA19 panel of assessors were also on the equivalent RAE2008 panel (see earlier), we entertain the possibility that REF2014 assessors were equally susceptible to such an effect because they would have known broadly whether a university’s UoA had a high or low reputation. Consistent with Hatch and Yanow (2003), we take the view that prior knowledge of any phenomenon filters and shapes what people then see and understand. Theory free observation is next to impossible. Therefore it is difficult to see how long-standing reputational impressions could be set aside when assessors considered ‘Environment’.

Also pertinent is what Coombs and Holladay (2001, p.338) describe as the ‘velcro effect.’ This arises where a negative history ‘sticks’ to an organization (such as a university, or equally a UoA). Thorndike (1920) was well aware of this possibility. In developing his halo theory, he observed that soldiers who had an inferior overall rating were also more likely to be judged poorly on specific criteria – what he called the ‘Devil Effect’. Much subsequent research has established that ‘bad’ experiences or impressions influence judgements much more than those deemed to be ‘good’ (Baumeister *et al.*, 2001). This suggests that once the ‘velcro’ or ‘Devil’ effects take hold they may be hard to escape, since more ‘good’ information is needed to compensate for the bad. Thus, reputation may be seen as an imaginative ‘iron cage’ that disciplines the observations and judgements of evaluation panels, despite their efforts to take an objective view.

Unlike ‘Impact’ case studies (where corroborating evidence was expected, and some auditing was undertaken), little supporting evidence was required in ‘Environment’ submissions. Where there was an absence of evidence in ‘Impact’ submissions, this ‘meant that ... quality of writing had a large effect’ (Manville *et al.*, 2015, p.xiv). Given that ‘Environment’ narratives were longer than ‘Impact’ narratives, there is potential for writing quality to have an even larger effect in ‘Environment’ submissions, and for HEIs to use language-related techniques to manage their image.

REF2014 panellists were instructed to assess submissions based on the ‘vitality’ and ‘sustainability’ of the research ‘Environment’ (REF, 2012, p.77). The Chair and Deputy-chair of the B&M panel have described how assessors awarded high scores to statements that ‘evidence vitality’ and ‘sustainable plans’ (Pidd and Broadbent, 2015, p.9). These instructions were conveyed to HEIs at advisory workshops. They provided strong incentives for submissions to attempt to influence assessors by creating a good impression through language

choice: for example, by being more appealing, more coherent, more ‘reader friendly’, and by asserting that their research environment was characterised by ‘vitality’ and ‘sustainability.’ If so, this has implications for understanding the extent to which gaming the system has (or can) become central to the whole exercise.

Research method

We analysed 98 of the 101 submissions regarding ‘Environment’ in UoA 19. These were downloaded from <http://results.ref.ac.uk/Results>. The number of FTE staff submitted by each university determined permissible length. Seven pages were allowed for between 1 and 14.99 FTE. A further page was allowed for each additional 10 FTE up to 54.99 FTE, and each additional 20 FTE beyond that. The longest submissions were by Lancaster and Manchester (122 FTE each: 15 pages permitted). The average length of submission was 5,770 words, ranging from 3,453 words (York St John) to 9,767 words (Manchester). HEIs were required to supply an overview and sections reporting on research strategy; people; income, infrastructure and facilities; and collaboration and contribution to the discipline or research base (REF, 2012). HEIs could determine the length of any section (subject to their overall length limit).

Instructions to assessors reinforce the appropriateness of using an impression management lens. They were urged to apply two generic but vague criteria: ‘vitality’ and ‘sustainability.’⁴ They ‘looked for clear evidence that [the submission was] feasible, well-considered and convincing’ and that the UoA sounded ‘like a great place to work, in which senior and junior researchers should thrive’ (Pidd and Broadbent, 2015, p.7). Despite this, there were no requirements to submit evidence of staff satisfaction or staff turnover. Assessors were directed to use quantitative information solely as a ‘crude indicator of overall activity’ because interpreting such statistics on a per capita/FTE basis ‘was difficult and probably not meaningful’ (Pidd and Broadbent, 2015, p.8, p.11).

The results of assessments were reported simply: for example, for London Business School, 75 per cent of the submission was graded at 4*, 12.5 per cent at 3*, and 12.5 per cent at 2*. Results were summarised widely as a ‘Grade Point Average’ (GPA). In the present example, the GPA is 3.625 ($[4 \times 0.75] + [3 \times 0.125] + [2 \times 0.125]$).

We ranked submissions according to their GPA score for ‘Environment’ (from best to worst). To resolve deadlocks from equal GPA scores, the higher(est) rank was assigned to HEIs submitting the larger(est) number of FTEs. Submissions of the top five HEIs were all scored at

the 4* level (GPA = 4). They were ranked (in descending order), based on the FTE submitted (in parentheses), as follows: Lancaster (122), LSE (81), Cardiff (73) and Strathclyde (73), Cambridge (39). The lowest GPA of 0.625 was for York St John (7). Submissions were divided into quartile groups (designated Q1, Q2, Q3, Q4) based on the above ranking scheme (see Table 1). Q1 comprises the 24 highest ranked HEIs. Table 1 reports GPA for HEIs, whether a member of the Russell Group, and whether any staff served as a UoA19 assessor.

Table 1: ‘Environment’ Quartile Groups
(by GPA score and then FTE)

Quartile 1 n = 24 Mean FTE = 63.8		Quartile 2 n = 25 Mean FTE = 36.2		Quartile 3 n = 25 Mean FTE = 24.2		Quartile 4 n = 24 Mean FTE = 11.0	
GPA	Institution (FTE)	GPA	Institution (FTE)	GPA	Institution (FTE)	GPA	Institution (FTE)
4	Lancaster (122)*	3.25	<i>Edinburgh (52)*</i>	2.75	<i>Glasgow (40)</i>	2.125	Sch Orient Asian (23)
	<i>LSE (81)</i>	3.125	<i>Liverpool (45)</i>		Bangor (29)		Bedfordshire (14)
	<i>Cardiff (73)*</i>		Sussex (44)		<i>Bristol (28)</i>	2	Leeds Beckett (17)
	Strathclyde (73)*		Stirling (43)		East Anglia (24)		Lincoln (9)
	<i>Cambridge (39)*</i>		Middlesex (40)		Ulster (22)	1.875	Westminster (21)
3.875	<i>Manchester(122)*</i>		Herriot Watt (37)		Bradford (19)		Hertfordshire (14)
	Bath (65)*		Swansea (28)*	2.625	Kent (43)		Northampton (12)
	<i>Imperial Lon.(58)</i>	3	Brunel (61)		Plymouth (33)		South Bank (9)
	<i>Oxford (42)</i>		Essex (50)		Huddersfield (19)	1.75	Greenwich (29)
3.75	<i>Leeds (73)*</i>		Kingston (25)*	2.5	<i>Newcastle (60)</i>		Dundee (9)
	Aston (46)*		Aberdeen (13)		<i>Queen Mary (33)*</i>		Birmingham City (5)
3.625	London Bus S (99)	2.875	Leicester (60)		Northumbria (23)		East London (3.25)
	Cranfield (41)*		<i>Queens Belfast (54)</i>		Bournemouth (21)	1.625	Glasgow Cale. (15)
3.5	<i>Nottingham (89)*</i>		<i>Exeter (49)*</i>		Open (18)		Anglia Ruskin (14)
	Reading (40)*		Roy Holloway (42)*		Staffordshire (7)		Derby (11)
	St. Andrews (22)*		Portsmouth (41)	2.375	Keele (18)		Teesside (6)
3.375	<i>Durham (45)</i>		Manch. Met. (26)		Aberystwyth (17)		Roehampton (5)
	<i>Sheffield (35)</i>		<i>York (23)</i>		Cent Lancs (11)	1.5	Sheffield Hallam (7)
	<i>Southampton (34)</i>		De Montfort (22)		Wolverhampton(11)		Chester (6)
3.25	City Uni Lond (78)		Coventry (17)		Edin. Napier (10)	1.375	Worcester (9)
	<i>Birmingham (53)*</i>		Salford (17)	2.25	West England (34)		London Met (4)
	<i>Kings-London (36)</i>		Brighton (16)		Birkbeck (30)	1.25	West Scotland (11)
3.5	Loughboro (61)*		<i>Uni Coll London (13)</i>		Notts Trent (23)	0.875	Sunderland (5)
3.25	<i>Warwick (104)*</i>	2.75	Hull (44)*		Robert Gordon (7)	0.625	York St. John (7)
			Surrey (42)	2.125	Oxford Brookes (24)		

* = supplied an assessor to the UoA19 B & M panel. Italic font = Russell Group member.

Fourteen of the 24 Russell Group universities are in Quartile 1, six in Q2, and four (Glasgow, Bristol, Newcastle, Queen Mary) are in Q3. No Q4 university (and only Queen Mary in Q3) supplied an assessor. Fifteen of the 24 universities in Q1 did so. We analysed the submissions using the four methods described below.

Word choice

We compiled lists of synonyms for the two major assessment criteria of ‘vitality’ and ‘sustainability’ using *Thesaurus.com*. This yielded 44 words (or word stems) for ‘vitality’ and 11 for ‘sustainability’ (see Table 2). To accommodate differences in submission length, frequency counts were standardised by calibrating frequencies per 1000 words. We then compared the frequency of all synonyms across quartile groups. Our expectation was that the frequency of synonyms for ‘vitality’ and ‘sustainability’ would be significantly higher for Q1 than Q4.

Table 2: Synonyms for Vitality and Sustainability

Assessment Criterion	Words or Word Stems
Vitality	clout; continuity; endurance; exuberan*; intensi*; spunk; stamina; strength*; verve; vigour*; animat*; ardour; audaci*; bang; being; bloom*; bounc*; driv*; existence; fervour; force*; get-up-and-go; go; guts; life; liveli*; lustiness; pep; pizzazz; power*; puls*; punch*; robust*; snap; sparkl*; starch; steam; stuff; venturesome*; vim; vivaci*; zest; zing; zip
Sustainability	continu*; viab*; feasib*; unceasing; green; imperishable; livable; renew*; support*; unending; worthwhile

Style characteristics

We used the ‘grammar and style check’ of Microsoft Word to determine the incidence of passive voice and infelicities of style, such as ‘long sentence’ and ‘wordiness’, and grammatical errors. Our view was that these infelicities would impair readability and negatively affect assessor disposition. We defined an ‘incoherence index’ as the sum of the frequencies of these infelicities per 1000 words. We expected that Q1 universities would demonstrate a lower incoherence score than Q4 universities. Additionally, we explored the incidence of cliché and jargon. Our expectation was that Q1 universities would have a greater ‘sensitivity’ to the proprieties of language use and a lower frequency of clichés and jargon than for Q4 universities.

Keywords

Keywords were identified using *WordSmith Tools 6* (Scott, 2012). A log-likelihood calculation identified keywords that occurred significantly more often in Q1 submissions (‘positive’

keywords). We compared these with keywords that occurred significantly more often in Q4 submissions ('negative' keywords).

Close readings

We selected four submissions as case studies and read them closely to assess whether research strategies were ambitious, feasible and clearly articulated; whether staff development was a priority and linked to underlying research strategy; and whether researcher support mechanisms were described clearly. The close readings were intended to reinforce or contradict findings reported elsewhere — or to otherwise illuminate the research question. However, we draw attention to the contestable nature of close reading commentaries. Close readers have limited capacity to deal with a 'plurality of plausible explanations' (Ron, 2008, p. 291) that exist for the complex array of social and organizational matters they canvass, and to enter interpretations in an unbiased fashion. Thus, the commentaries we make should not be viewed necessarily as more definitive than other explanations.

Cardiff (GPA = 4, FTE = 73, Q1) was selected because of its keenness to improve its ranking in performance league tables⁵. Swansea (GPA= 3.125, FTE = 28, Q2) was chosen to explore how it could be assessed as 'a great [or even tolerable] place to work' given that 25 teaching staff resigned after a new Dean (Nigel Piercy, appointed July 2013) introduced controversial changes without consulting staff. Piercy's reign as Dean is alleged to have been 'toxic'; to feature an 'abrasive management style'; and to have engaged in 'gratuitously offensive' and 'puerile' diatribe against staff and students including that they were 'unpleasant and grubby little people' (<http://waterfrontonline.co.uk/news/university-faces-tough-questions-as-piercy-resigns>). Greenwich (GPA=1.75, FTE = 29, Q4) was chosen because it had an almost identical number of staff as Swansea, but fell in Q4, despite the absence of a 'toxic' management regime. Sunderland (GPA= 0.875, FTE = 5, Q4) was chosen to enable exploration of the characteristics of one of the lowest-scored submissions.⁶

Results

Synonyms for 'vitality' and 'sustainability'

The four most frequent synonyms for 'vitality' were 'strength*' (n = 594), 'driv*' (n = 157), 'force*' (n = 157), and 'vital*' (n = 82).

Table 3: Frequency of Synonyms for Vitality and Sustainability

Quartile	Total Words	Average Words per Submission	Vitality Synonyms	Vitality Synonyms per 1000 words	Sustainability Synonyms	Sustainability Synonyms per 1000 words	Sustainability and Vitality Synonyms per 1000 words
1	175,925	7330	470	2.67	1226	6.97	9.64
2	149,082	5963	482	3.23	1042	6.99	10.22
3	130,320	5213	302	2.32	954	7.32	9.64
4	110,226	4493	290	2.63	760	6.89	9.52
Total	565,553	22999	1544		3982		
Mean	141,388	5750	386	2.73	996	7.04	9.77

For Q1 and Q4, the average frequency per 1000 words of synonyms for vitality and sustainability are almost identical (vitality: 2.67 vs 2.63; sustainability: 6.97 vs 6.89). This indicates no obvious association between words conventionally connected to vitality and sustainability, and reported ratings of submissions. Given that Pidd and Broadbent (2015, p.9) acknowledge high scores were awarded to statements that ‘evidence vitality and ... sustainable plans’, this suggests assessor judgements may have rewarded broader narrative style rather than be deceived by word choice, or have been influenced by their pre-conceptions of each HEI. A process of isomorphism seems evident. All universities seem likely to be aware of the need to employ linguistic tropes to signal vitality and sustainability. The situation could hardly be otherwise, given that this was so clearly signalled. The submissions of HEIs therefore coalesce around common linguistic forms for these issues, despite their actual practice differing substantially.

Style characteristics

Table 4 reveals that submissions in Q1 were 30% less likely to use the passive voice than submissions in Q4 (5.87 vs 8.34 per 1000 words). A similar result was obtained for the ‘incoherence index’ (the rate of observed incoherence factors per 1000 words). The index value was 61% less in Q1 than in Q4 (3.29 vs 8.34).

Table 4: Passive Voice and Other Style Characteristics

Quartile	Total Words	Use of Passive	Passive per 1000 words	Incoherence Count*	Incoherence Index**	Clichés, Colloquialisms & Jargon	Clichés, Colloquialisms
----------	-------------	----------------	------------------------	--------------------	---------------------	----------------------------------	-------------------------

							& Jargon per 1000 words
Q1	175,925	1033	5.87	578	3.29	75	0.43
Q2	149,082	993	6.66	497	3.33	44	0.30
Q3	130,320	959	7.36	728	5.59	34	0.26
Q4	110,226	919	8.34	629	8.34	31	0.28
Total	565,553	3904		2432		184	
Mean	141,388	976	6.90	608	4.30	46	0.33

* Defined as overlong sentences, ‘wordiness’ and/or fragments of correct sentences that should not stand alone (e.g., because there is no subject or no verb)

** Incoherence count per 1000 words

Analysis of the clichés and jargon identified by *Microsoft Word* reveals an average incidence of 0.28 per 1000 words for Q2, Q3 and Q4. The frequency for Q1 (0.43) was 54% higher, suggesting that assessors perceived some clichés and jargon positively. We therefore constructed a set of ‘*superiority*’ clichés and jargon (see Table 5). These were considered likely to be used to assert a university’s superiority and research excellence (e.g., expressions such as ‘cutting edge’).⁷

Table 5: ‘Superiority’ Clichés and Jargon by Quartile
(* and derivative words)

	Q1	Q2	Q3	Q4	Total
Words	175,925	149,082	130,320	110,128	565,553
Track(-)record	13	11	15	10	49
Impact*	336	257	220	178	991
World(-)class	32	26	11	8	77
Showcase*	12	5	5	9	31
Cutting edge	7	11	4	0	22
Research excellence	36	27	16	24	103
grow*	172	165	149	132	618
award	464	313	290	157	1224
prize	68	38	17	9	132
Totals	1140	853	727	527	3247
Incidence per 1000 words	6.48	5.72	5.58	4.79	5.74

There was a 35% higher level of ‘superiority’ clichés and jargon per 1000 words in Q1 than in Q4 (6.48 vs 4.79). ‘Prize’ and ‘world class’ were used 370% and 300% respectively more frequently in Q1 than in Q4. Tables 4 and 5 reveal that Q1 submissions were less prone to incoherence (as defined here) and the use of passive voice. They were much more likely than Q4 submissions to trumpet superiority and excellence through clichés and jargon.⁸

Keywords

Results were separated into the eight themes shown in the left hand column of Table 6.

Table 6: Keywords

Theme	Positive	Negative
Self-reference and university structure	DUBS, SBS, LUBS, LUMS, BBS, Imperial, Cardiff, SBE, faculty, centres, DOF, DOA, SMS, DOM, LSE, school	Unit, university, UOA
Subject areas	Science, financial, risk, finance, behavioural	Communications
Funding	ESRC, EPSRC, DTC, NIHR, fellowships, REF	
Staff	Walker, Taylor, scholars, tenure, associate	Reviewer, invited, appointed, Sun, professor, leader, staff, Prof, Dr
Students		Completions, supervisory, students, postgraduate, doctorate, student
General positive	Top, major	Experienced, increase, active
other	corporate, agendas, innovation,	Conference, submitting, rural, China, Int, development, outputs, enterprise, research
Grammar words	's	She, he, upon, will, to, within

One striking result is that 16 ‘self-reference’ positive keywords were used in Q1 compared to three (generic) negative keywords in Q4. The reluctance of Q4 institutions to use specific self-reference keywords seems to reflect tacit acknowledgement of the lower power of their ‘reputational brand.’ Higher ranked institutions had an opposite perception of their brand value. ‘Funding’ terms linked to the UK Research Councils (*ESRC, EPSRC, DTC, NIHR*) stand out on the list of positive keywords, signalling superiority and quality. In contrast, no ‘funding’ keywords were significant in Q4 submissions. ‘General positive’ keywords reinforce the notion of supremacy: Q1 institutions used ‘top’ or ‘major’ while Q4 institutions used ‘experienced’ or ‘active.’ This is consistent with the increased frequency of superiority clichés (such as ‘cutting edge’) in Q1 submissions.

‘Staff’ keyword differences are less easy to interpret. Several simply reflect common names (at least three staff members named ‘Walker’ are referred to by Cardiff). The paucity of Professors (and even Doctors) among several Q4 submissions is reflected in a desire to acknowledge such titles. This practice is absent in the majority of Q1 submissions. Similarly,

Q4 submissions were keener to state that staff had been ‘appointed’, ‘invited’ or were a ‘reviewer.’ Among better-ranked HEIs, such activities largely go unmentioned, since they are regarded as the norm.

The emphasis in many Q4 submissions on terms relating to student research (*completions*, *doctorate*) seems to have drawn attention to the relationship between FTEs submitted and PhD completions. Pidd and Broadbent (2015, p.8) affirmed that the ‘panel was concerned that some submissions included far too many PhD enrolments for the number of staff included’, and that such concerns ‘typically led to a lower score.’ This affirmation is corroborated by our finding of no statistically significant link between GPA score and PGR completions per FTE. Manchester (Q1) reported 2.64 completions per FTE staff submitted (n = 122). In contrast, the University of South Wales (Q4) reported 13 completions per FTE (n = 3).

An apostrophe followed by ‘s’ (that is, ‘’s’), was a positive keyword, used commonly in phrases such as ‘the faculty’s research.’ Such use emphasises ownership of the research agenda by an organizational unit rather than an individual, and fosters a perception of inclusivity. The negative keyword ‘will’ is often used to refer to plans and activities that are presently unfulfilled – an interpretation supported by our close readings.

Overall, the positive keywords in Q1 submissions are consistent with a ‘finished article’ discourse. This is unsurprising given that these institutions were generally long-established. The mean age of Q1 universities was 168 years, whereas for Q4 universities it was 26 years. Q4 submissions had a higher level of coyness and more of a ‘we are developing’ discourse.

Close reading

Overview. Cardiff’s submission is characterised by a tone of active, forceful and forward-thinking confidence. Swansea highlights a ‘consolidation of research’ that has taken place ‘under the leadership of a new Dean.’ The submissions of Greenwich and Sunderland tend to offer aspirational mantras in lieu of hard data.

Greenwich gives the impression of an uphill struggle to establish a research culture, with stringent oversight necessary to guard against shirking. Its stress on systems and ‘monitoring’ is notably absent in submissions of long-established universities. Greenwich conveys weakness rather than strength. Sunderland begins with an obvious statement of weakness: of not entering the RAE in 2008 because of ‘significant structural and managerial change for the faculty...’ Its submission is high on ambition, but vague on details. Such a mixture seems unlikely to

convince readers that organizational structures and strategies are in place to deliver the ambitious outcomes mentioned.

Research strategy. Cardiff claims that it ‘aims to attain research excellence in breadth and depth ... [and that] ... research lies at the heart of the School’s mission and strategy.’ This claim is supported by details of ‘research outputs’, PhDs awarded, and prize-winning students. Cardiff identifies five key features of its research and provides convincing supporting detail, emphasising ongoing investment in recruitment, and links with university-wide research centres. Simultaneously, it highlights operational procedures (e.g., involving the doctoral programme).

Swansea affirms the ‘School’s vision is to be a research-led, internationally focused centre of excellence capable of supporting and sustaining research of the highest calibre.’ This ‘vision’ will be achieved through promoting a ‘strong, collegiate research environment’ and committing strategically to ‘attract and retain talented research-active staff and research students.’ Swansea is aspirational in declaring it will deliver further increases in publications in leading journals, higher citation rates and more research funding. But it does not specify how these increases will be achieved — in contrast to Cardiff. It vaguely affirms that making ‘internationalization central to the School’s research strategy’ will be achieved by commitment to ‘invest in new strategic partnerships in ... India, China and Africa.’ One wonders whether such ‘aspirational’ (but imprecise) commitments carried much weight with assessors.

Greenwich’s claims seem feasible, but not ambitious: for example, the pledge to increase ‘the proportion of staff research active (the publication of at least one output at 1* or above each year) to 75% by 2017.’ Much is made of the prospect that research groups will achieve publication ambitions and secure increased research and enterprise income of 10% annually. There is repeated reference to staff research activity being ‘monitored’ and to research active staff having ‘to account for how they use that time.’ This implies that staff will shirk their research responsibilities without monitoring. It may conflict with the long-standing research finding that academics tend to have a particularly strong need for autonomy, independence and individual expression (e.g. Lindholm, 2004), and so undermine a fundamental aspect of what most academics would see as ‘a great place to work.’

Sunderland claims its vision ‘is that all academic staff ... will become actively involved in research to develop a thriving and sustainable research culture and an internationally

renowned research centre.’ This is a bold ambition for a school that submitted only five of 38 FTE. Although Sunderland outlines an aggressive recruitment strategy and a commitment to develop three existing research centres, the contrast between its grand vision and current position strains credulity. Offering bold ambitions that depart heavily from what exists at present risks these offerings being read as bombast.

Research people. Cardiff does not cite overall staffing figures, but states that 83% of academic staff are appointed on open-ended teaching and research contracts — a major signal of research intent (even if only 56% of eligible staff were submitted).⁹ Cardiff claims to have experienced low staff turnover (without specifying the levels or basis of comparison). A clear and comprehensive outline is provided of staff development plans and staff support. The operational detail offered supports the more general statements of strategic intent.

Swansea stresses the importance of the Dean in effecting a ‘transformation’ that it claims ‘has brought renewed energy and ambition to the School, establishing enhanced research performance and impact as a key strategic priority.’ One wonders whether ‘transformation’ is a euphemism to disguise a significant level of staff resignations — and to signal the recruitment of new staff more sympathetic to the Dean’s ‘vision.’ Swansea highlights how research is supported (and refers to a ‘collegial environment’). However, none of the pains associated with its ‘transformation’ are acknowledged. This accords with the contention that ‘Environment’ statements encourage authors to conceal weaknesses where possible, despite the likelihood they will be known to assessors. Swansea was well-rated (equal 26th) despite widespread public knowledge of its ‘toxic’ workplace environment, contrary to the ‘velcro’ effect.

Greenwich was probably unwise to signal that the ‘inevitable result of success’ was a staff exodus. The ‘bad is stronger than good’ effect suggests that it would then take many more positive achievements to offset the effects of this information (Baumeister *et al.*, 2001). If Greenwich really was ‘such a great place to work’, why would successful staff leave? It is equally unwise, from the standpoint of impression management, to signal spending ‘a considerable amount on the development of academic staff (£61.6K in 2012/13)’ when the amount is barely £2,000 per FTE submitted. Sunderland indicates upheaval and a ‘number of structural changes’ in 2012/13, and a ‘significant development’ of research staff since 2011. However, it fails to elaborate. Sunderland emphasises the need to recruit staff with doctorates.

It provides biographies of the staff submitted to REF2014 – something not requested in the assessment criteria (REF2012, p.75) – and therefore information of dubious value.

Research income, infrastructure and facilities. Cardiff highlights its attainment of £10.4 million in research income, placing it in the top quartile of Russell Group universities for total income and average income per FTE. It is clear that research income generation is an important priority.

Swansea's major successes are highlighted (e.g., PhD programme growth). However, this is the only 'Environment' submission that credits a named Dean explicitly for effecting a 'transformation' (implying a major advance on an unsatisfactory past). This suggests a top-down managerialist approach. It seems inadvisable to associate claims of progress so closely with one senior manager, particularly one whose appointment is recent. This would appear to torpedo any claim that a successful and sustainable research environment existed before his arrival. Moreover, the Dean left the university by mutual agreement in July 2015. The Swansea submission exaggerated his achievements and attempted to conceal the problems his tenure created.

Greenwich highlights three examples of staff successes in income generation but does not disclose actual figures. Given the entire section comprises barely half a page (353 words compared to Swansea's 829 words), the impression is that Greenwich did not have much to offer researchers in terms of income, infrastructure and facilities. Sunderland presents a relatively lengthy discussion of research students and describes the support they receive to 'ensure that their work is of the appropriate standard.' No hard data are offered on numbers, enrolments, or completion rates. No details of research income are provided, suggesting there was none to report.

Collaboration and contribution. Cardiff details many contributions of staff to the discipline – through editorships, membership of editorial boards, participation in academic and professional bodies, and high-profile research collaborations. Swansea adopts a similar strategy, but offers fewer examples. Greenwich emphasises that 'staff participate', 'staff collaborate', and 'staff support.' However, it does not elaborate, apart from naming five Visiting Professors and Fellows. Sunderland replicates sentences used in its 'People' sub-section and lists the journal publications of each of its five submitted staff members. However, the quality of the journals

cited does not inspire confidence that Sunderland can transition to an ‘internationally renowned research centre’.

Summary. The highest ranked submission, Cardiff, is distinguished by its sense of activity, and the specific steps it identifies to realise strategy. Although Cardiff offers a compelling story, questions should be raised about whether a Business School which does not include a large proportion of its eligible staff in REF2014 is really such a ‘great place to work’. Impression management is evident in the lack of explicit consideration of ‘problems’ by Cardiff and Swansea — even when manifestly obvious. Swansea puts a positive spin on high levels of resignations, describing them euphemistically as a ‘transformation’ in staffing. Greenwich and Sunderland attempted to hide shortcomings (e.g. Sunderland’s low Ph.D completions, Greenwich’s low research income). Thus, the submission writers seem keen to conceal whatever difficulties they can, and to bluster their way through the rest. It is difficult to accept that submission narratives can convey an accurate picture of whether an institution is ‘a great place to work.’ They are more akin to a process of self-certification that is distorted by obfuscation and systematic exaggeration of achievements.

Discussion

The impression management potential afforded to HEIs in research evaluation exercises (and specifically in the ‘Environment’ component of REF2014) has been largely ignored. This is surprising given the financial benefits ensuing.¹⁰ Lancaster (122.38 FTE) scored a GPA of 4 under ‘Environment’, triggering a payment of almost £620,000 in the academic year 2015/6 for this element of REF2014. Cardiff’s selective approach (72.6 FTE) still delivered an ‘Environment’ return of about £360,000 in 2015/6. Swansea submitted 27.9 FTE, received a GPA of 3.125, and were rewarded with just over £57,000. In contrast, Greenwich submitted 0.8 more FTE (28.7) than Swansea, obtained a GPA of 1.75, and received only about £4,500. The rewards for higher-ranked ‘Environment’ submissions are clearly substantial.

There are some distinctively different language-related characteristics between high-ranked submissions and low-ranked submissions. Higher ranked institutions have:

- a much lower incidence of passive voice
- a much lower index of incoherence

- a much stronger use of ‘superiority’ clichés and self-referential keywords
- a ‘finished article’ discourse rather than a ‘we are developing discourse’; and
- a tendency to cite specifics rather than generalities to support arguments.

We found no differences between high-ranked and low-ranked universities in terms of frequency of synonyms for ‘vitality’ and ‘sustainability, despite clear instructions to the assessing panel in this regard. Our supposition is that since the need for such synonyms was so clearly communicated, they were widely employed across the sector, irrespective of whether they bore much relationship to reality.

Policy implications

The HEFCE’s *Consultation on the Second Research Excellence Framework* (2016) embodies many of the recommendations of the Stern Review (2016) regarding how future research exercises are to be shaped. The desire for ‘continuity with REF2014’ (HEFCE, 2016, note 10) promises a ‘more structured template’ which decreases ‘... narrative elements of the template and increas[es] the use of data which is already collected and held by institutions’ (note 112). This seems to be a tacit admission that narrative has induced unintended bias.

We concur with the Stern Review (2016, p.23) that ‘we should reward those institutions which have a dynamic and creative research environment [and] a vision and direction for their research and related activities.’ The difficulty lies in identifying objectively such oases of excellence. Metrics can help. The HEFCE decision (2016, note 113) to work with the *Forum for Responsible Metrics* in developing appropriate indicators in this regard is commendable. But what form might such metrics take? Few seem likely to reject the proposition that there is a strong and positive correlation between a dynamic research environment and the number and percentage of research active staff (FTE) employed by a submitting UoA.¹¹ Equally, research income (ideally per FTE) can indicate a strong research environment that provides time (through teaching buy-out) and resources to enable academics to develop their ideas. However, excessive attention to achieving grant income targets may under-value levels of support that are critical for academics to perform their research.

Given that assessments of research ‘Environment’ also consider ‘how good’ a place is to work, we suggest that levels of staff turnover be considered. Excessive staff turnover is an indicator of discontent and a useful proxy for the quality of a work environment (Mobley,

1979). However, it has not featured directly in evaluation of research environment. Nor have indicators of how staff feel about their work environments. This is despite mounting evidence that a growing culture of audit, targets and rankings has generated immense pressure on academics, perhaps particularly within business schools (Craig et al., 2014). It has produced research environments in which journal of publication is more important than content (Butler and Spoelstra, 2017). The conclusion of Gabriel (2010, p.769) is apt: 'I doubt that there are many professions whose members are so relentlessly subjected to measurement, criticism and rejection as academics, exposing them to deep insecurities regarding their worth, their identity and their standing.' Such dynamics do not help the production of useful, meaningful ideas-driven research (Alvesson et al., 2017). We therefore suggest incorporating measures of staff satisfaction into assessments of research environment. Do academics feel supported in their research? Or do they feel harassed, stressed and over-burdened by a proliferation of performance targets? These are critical issues, and it is sensible and feasible to measure them. Preferably, any staff satisfaction survey would be conducted by an independent authority, such as HECFE, and not by institutions themselves – and be at the UoA level rather than the university level to avoid disadvantaging well-run UoAs in poorly-run universities. As is the case with surveys of student satisfaction (e.g. the National Student Survey in the UK) there is an obvious risk of game playing. Survey respondents could be pressured to report more positive attitudes than they feel to avoid damaging an institution's brand. The steps taken to minimise this risk with students could be adopted here as well.

The case for considering PGR completions (or completions per FTE) is less clear. While PGR students are an important part of a research environment, over-recruitment leads to onerous supervisory burdens and poorer supervision. Panels have treated such data with caution, given the exclusivity of recent research assessment exercises.¹² This may account for why we found 'PGR completions' was not significant in the regression analysis. However, if all research active staff are entered in assessment exercises, 'PGR per FTE submitted' will gain credence as a measure of research environment. If so, it is vital to avoid crudely emphasising volume, since this will provide institutions with incentives to recruit PGR students merely to achieve what can easily become high volume targets. While low PGR numbers are an important indicator, high PGR to staff ratios are also an indicator of a poor research environment, and should be considered critically.

A further option is for the next REF framework to develop new metrics such as ‘research concentration’ ratios, perhaps based on the Herfindahl-Hirschman Index.¹³ Consider, for example, the current proposal that requires all research active staff to be submitted, with an average of two outputs and a minimum requirement of zero for each FTE staff member (HEFCE, 2016, questions 8 and 9a/c). Consequently, if HEI A and HEI B find themselves each obliged to submit 60 staff, this will require the reporting of 120 outputs each. Assume HEI A chooses to submit 20 staff with six outputs each, and the remaining 40 staff with zero outputs. Assume HEI B chooses to submit all 60 staff with two outputs apiece. Which is the better research environment? We submit that most would concur that it is HEI B.

Thus, a further issue is how many publications staff should be required to submit in future REF-type exercises. Stern (2016) has recommended that all staff now be included, with a minimum of two publications but a maximum of six. However, this will not remove the problems discussed here. All staff may now be included, but only as a multiplier to determine the total number of outputs to be submitted. Institutions are still likely to evaluate outputs, perhaps continuing to rely heavily on the Chartered Association of Business Schools’ *Academic Journal Guide* – a practice widely criticised (e.g. Tourish and Willmott, 2015). Game playing would remain, and consume precious institutional resources. For example, a UoA may decide that Dr Y should be submitted with six publications deemed to be 4*, but Dr X should be submitted with none, since these are deemed to be 3 or below. In this scenario, individual academics could be rewarded handsomely for producing up to six 4* papers. Those that ‘fail’ to do so could still be penalised, for example, by being moved to teaching-only contracts. Complex systems for scrutinising outputs would remain. The ‘Environment’ would suffer from divisions between academic colleagues and the alienating effects of performance measurement systems.

This problem could be ameliorated if the average number of outputs required remained at two, but the maximum was reduced from six to four. An additional benefit is that while academics would still focus much of their efforts on the needs of the REF, they would also be freer to pursue a research agenda driven more by their own intrinsic interests – for example, by publishing papers that required longer gestation times, and even publishing books. Our suggestion here would help to prevent REF research ‘Environment’ exercises remaining an obsession that, paradoxically, harms the research environment it is attempting to evaluate.

We should not place excess faith in metric-driven analyses of ‘Environment’ and highly-structured templates. The latter will risk compressing heterogeneous activities, strengths and weaknesses, into bland and rigid snapshots of the research ‘Environment.’. Unintentionally, such analyses can disguise actual variations in institutions. We need to draw out the differences between institutions, and explore more deeply the ‘Environment’ narratives that are expected to be disclosed. To do this, it would be beneficial to specify how we conceive an ideal research ‘Environment’. Should such an ‘Environment’ be conceived as having desirable features other than those touted currently? Should they include the promotion of creativity and the tolerance of risk-taking?

The way ahead

A variety of strategies is being deployed in narratives to inflate successes and downplay problems. HEIs with an established reputation are attempting to take advantage of the halo effect – that is, capitalise on assessors’ existing knowledge of their positive position, by playing to their preconceptions. Those with poor reputations attempt to escape the velcro effect – a more difficult task given the exaggerated potency of negative information in shaping perceptions. Swansea scored well in its Environment submission, despite a growing reputation for aggressive management by its Dean. The assessor panel evidently concluded that other strengths compensated for this problem.

The findings we report should attract the attention of policy makers who are dedicated to ensuring that future research assessment is ‘fit for purpose’ and that ‘research funding is allocated more efficiently’ (Stern Review, 2016, p.3)]. If the findings are generalizable to a broader catchment of UoAs (as we believe)¹⁴ this should prompt stronger effort to devise better ways of assessing research ‘Environment’ across all disciplinary boundaries.

The ambient research environment in HEIs is critical to fostering research of service to society. However, attempts to evaluate ‘Environment’ seem prone to distort what they are trying to evaluate. There is a strong argument for assessment of ‘Environment’ to be refocused, given widespread acceptance that ‘publicly-funded universities should be accountable for what they do and for whether they provide graduates and research of service to society’ (Craig et al., 2014, p.2). It should be a matter of considerable public debate whether public funds are expended on future exercises conducted in the form of the REF2014 assessment of research ‘Environment.’ A strong case can be made that continued assessment of research

‘Environment’ (as in REF2014) will simply ‘homogenise’ universities and ‘irreparably harm [...] the creative paradoxes (see Marginson, 2010) that sustain the public university’ (Craig et al., 2014, p. 20). Further dialogue beyond that offered by Wilsdon et al. (2015) and Stern (2016) is needed to clarify how research in individual universities articulates with, and contributes to, the overall role of the public university system.

Appendix
Table 1: Multivariate Regression (OLS) Results

	Model 1		Model of Best Fit (Model 2)	
	‘Environment’ GPA score		‘Environment’ GPA score	
Variable	B (unstandardized beta coefficients)	Significance level	B (unstandardized beta coefficients)	Significance level
Constant	1.617**	.000	1.827**	.000
FTE	0.020**	.000	0.016**	.000
Income by FTE	0.000**	.001	0.000**	.004
Russell membership	0.242*	.050	0.206*	.095
Panel membership	0.346*	.014	0.375**	.008
Income	0.000*	.001		
PGR	0.002	.386		
PGR by FTE	-0.009	.900		
R ²	0.651		0.629	
F	23.98**		39.45**	

n = 98, * and ** indicate significance at the 10% and 1% levels respectively. Although ‘Income by FTE’ is significant, it has minimal impact on GPA score (complete coefficient value is 0.000001974 = 0.000 to three decimal places).

Model 2 is statistically acceptable. Relevant tests confirm it does not suffer from heteroscedasticity or multicollinearity (all Tolerances were above 0.6, and VIF below 2). Menard (1995) suggests that Tolerance should be above 0.2 and VIF below 10 to avoid multicollinearity problems. This was not the case in Model 1 (FTE has a Tolerance = 0.15 and VIF = 6.63; Income has a Tolerance = 0.13 and VIF 7.95, while PGR has a Tolerance = 0.15, VIF = 6.77).

Table 2: Descriptive statistics

Variable	Mean	Standard deviation	Min	Max
FTE	33.79	25.75	3.25	122.38
Income by FTE	90436.62	75216.73	270.46	402028.09
Russell membership	0.24	0.43	0	1
Panel membership	0.23	0.42	0	1
Income	3427068.21	4701978.50	1.76	25927917.00
PGS	48.56	44.49	0.50	320.95
PGS by FTE	1.56	0.94	0.04	5.00

References

- Alvesson, M., Y. Gabriel, and N. Paulsen (2017). *Return to Meaning: A Social Science with Something to Say*. Oxford: Oxford University Press.
- Amernic, J. and R. Craig (2007). 'Guidelines for CEO-speak: Editing the language of corporate leadership', *Strategy and Leadership*, **35**, pp.25-31.
- Baumeister, R., C. Finkenauer, and K. Vohs (2001). 'Bad is stronger than good', *Review of General Psychology*, **5**, pp. 323-370.
- BBC News. 2017. 'Universities challenged on top 1% advert', 8 June. Available at: <http://www.bbc.co.uk/news/education-40187452>
- BIS, 2016. *Lord Stern's Review of the Research Excellence Framework*. <https://www.gov.uk/...data/.../ind-16-1-ref-review-call-for-evidence.pdf>.
- Brass, K. and D. Rowe (2009). 'Knowledge limited: Public communication, risk and university media policy', *Continuum: Journal of Media and Cultural Studies*, **23**, pp.53-76.
- Butler, N., and S. Spoelstra (2017). 'How to become less excellent.' In T. Huzzard, M. Benner and D. Kararreman (Eds.), *The Corporatization of the Business School: Minerva Meets the Market*. London: Routledge, pp.74-91.
- CDBU (2015). 'A whole lotta cheating going on? REF stats revisited'. Retrieved from http://cdbu.org.uk/ref-stats-revisited/#_edn4.
- Cho, C.H., R.W. Roberts and D.M. Patten (2010). 'The language of US corporate environmental disclosure'. *Accounting, Organizations and Society*, **35**, pp.431-443.
- Coombs, W.T., and S.J. Holladay (2001). 'An extended examination of the crisis situation: A fusion of the relational management and symbolic approaches,' *Journal of Public Relations Research*, **13**, pp. 321-340.
- Elliot, J. (2010). 'How do past crises affect publics' perceptions of current events? An experiment testing corporate reputation during an adverse event.' In Coombs, W.T. and S.T. Holladay (Eds), *Handbook of Crisis Communication*, Wiley-Blackwell, pp. 205-220.
- Côté, J.E. and A.L. Allahar (2011). *Lowering Higher Education: The Rise of the Corporate Universities and the Fall of Liberal Education*. Toronto: University of Toronto Press.
- Craig, R., J. Amernic and D. Tourish (2014). 'Perverse audit culture and accountability of the modern public university', *Financial Accountability and Management*, **30**, pp.1-24.
- Derrick, G. and G.N. Samuel (2016). 'The evaluation scale: Exploring decisions about societal impact in peer review panels', *Minerva*, **54**, pp.75-97.

- Douglas, G. (2015). 'Law sub-panel generic feedback session'. Presentation to the Institute of Advanced Legal Studies, London, 4 March.
- Duarte, P., H. Alves and M. Raposo (2010). 'Understanding university image: A structural equation model approach', *International Review on Public and Non Profit Marketing*, **7**, pp.21-36.
- Forlie, E., L. Ashburner, L. Fitzgerald and A. Pettigrew (1996). *The New Public Management in Action*. Oxford: Oxford University Press.
- Forum for Responsible Metrics. (2017). *Metrics in REF2021: Advice from the UK Forum for Responsible Research Metrics*, <http://www.universitiesuk.ac.uk/policy-and-analysis/Pages/forum-for-responsible-research-metrics.aspx>
- Gabriel, Y. (2010). 'Organization Studies: A space for ideas, identities and agonies, *Organization Studies*, **31**, pp. 757-775.
- Geppert, J. and J.E. Lawrence (2008). 'Predicting firm reputation through content analysis of shareholders' letters', *Corporate Reputation Review*, **11**, pp.285–307.
- Geuna, P. and M. Piolatto (2016). 'Research assessment in the UK and Italy: Costly and difficult, but probably worth it (At least for a while)', *Research Policy*, **45**, pp.260-71.
- Hatch, M. and D. Yanow (2003). 'Organization theory as an interpretive science.' In H. Tsoukas and C. Knudsen (eds.) *The Oxford Handbook of Organization Theory: Meta-theoretical Perspectives*, Oxford: Oxford University Press, pp.63-87.
- HEFCE (2016). *Consultation on the Second Research Excellence Framework*, London: HEFCE.
- Kellard, N. and M. Sliwa (2016). 'Business and management impact assessment in Research Excellence Framework 2014: Analysis and reflection', *British Journal of Management*, DOI: 10.1111/1467-8551.12186.
- Larivière, V., Gingras, Y., Archambault, E. 2009. 'The decline in the concentration of citations, 1900-2007', *Journal of the American Society for Information Science and Technology*, **60**(4):858-62.
- Lindholm, J. (2004) 'Pathways to the professoriate: The role of self, others, and environment in shaping academic career aspirations', *The Journal of Higher Education*, **75**, pp. 603-635.
- Linkova, M. (2014). 'Unable to resist: Researcher's responses to research assessment in the Czech Republic', *Human Affairs*, **24**, pp.78-88.

- Manville, C., S. Guthrie, M-L Hensham, B. Garrod, S. Sousa, A. Kirtley, S. Castle-Clarke and T. Ling (2015). *Assessing impact submissions for REF: An evaluation*. Rand Corporation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1032/RAND_RR1032.sum.pdf.
- Marginson, S. (2010). 'The university: Punctuated by paradox', *Academic Matters* (May) pp.14–18.
- Matherly, L.L. (2012). 'A causal model predicting student intention to enrol moderated by university image: Using strategic management to create competitive advantage in higher education', *International Journal of Management in Education*, **6**, pp.38-55.
- Menard, S. (1995). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage.
- Mingers, J. and L. White (2015). 'Throwing out the baby with the bathwater: The undesirable effects of national research assessment exercises on research', *Arxiv.org*. doi:arxiv.org/abs/1502.00658.
- Mobley, W., R. Griffeth, H. Hand and B. Meglino (1979). Review and conceptual analysis of the employee turnover process, *Psychological Bulletin*, **86**, pp.493-522.
- Nisbett, R.E. and T.D. Wilson (1977). 'The halo effect: Evidence of unconscious alteration of judgments', *Journal of Personality and Social Psychology*, **35**, pp. 250-256.
- Owen.P.D., Ryan,M., Weatherston,C.R. (2007). 'Measuring competitive balance in professional team sports using the Herfindahl-Hirschman Index', *Review of Industrial Organisation*, **31**:289-302.
- Pidd, M. and J. Broadbent (2015). 'Business and Management Studies in the 2014 Research Excellence Framework', *British Journal of Management*, **26**, pp.569-581.
- REF2012. 'Panel criteria and working methods'. http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2C.pdf
- Rhoades,S.A. (1993). 'The Herfindahl-Hirschman Index', *Federal Reserve Bulletin*, **3**:188-9.
- Ron, A. (2008). 'Power: A pragmatist: Deliberative (and radical) view.' *Journal of Political Philosophy* **16**, pp.272–292.
- Roper, S. and G. Davies (2007). 'The corporate brand: Dealing with multiple stakeholders', *Journal of Marketing Management*, **23**, pp.75-90.
- Saunders, J., V. Wong and C. Saunders (2011). 'The research evaluation and globalization of business research', *British Journal of Management*, **22**, pp.401-19.

- Scott, M. (2012). *WordSmith Tools Version 6*, Stroud: Lexical Analysis Software.
- Slaughter, S. and G. Rhoades (2009). *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. Baltimore, MD: Johns Hopkins University Press.
- Stein, R.H. (1990). 'The issue of image', *Journal for Higher Education Management*, **6**, pp.3-6.
- Stern, N. (2016). *Building on Success and Learning from Experience: An Independent Review of the Research Excellence Framework led by Lord Stern*. Department for Business, Energy and Industrial Strategy. <https://www.gov.uk/government/publications/research-excellence-framework-review>.
- Taylor, J. (2011). 'The assessment of research quality in UK universities: Peer review or metrics?' *British Journal of Management*, **22**, pp.202-17.
- The Guardian. 2017. 'Advertising watchdog forces Reading University to ditch 'top 1%' claim', 8th June. Available at: <https://www.theguardian.com/education/2017/jun/08/advertising-watchdog-forces-reading-university-to-ditch-top-1-claim>
- Thorndike, E.L. (1920). 'A constant error in psychological ratings,' *Journal of Applied Psychology*, **4**, pp.25-29.
- Times Higher Education, (2015). 'REF cost almost £250 million'. 13 July.
- University of Cardiff, 2015. 'Quality and Performance'. <http://www.cardiff.ac.uk/research/impact-and-innovation/quality-and-performance>
- Tourish, D. and H. Willmott (2015). 'In defiance of folly: Journal rankings, mindless measures and the ABS guide,' *Critical Perspectives on Accounting*, **26**, pp.37-46.
- Wernick, A. (2006). 'Rebranding Harvard,' *Theory, Culture and Society*, **23**, pp.566-7.
- Wilsdon, J., L. Allen, E. Belfiore, P. Campbell, S. Curry, S. Hill, R. Jones, R. Kain, S. Kerridge, M. Thelwall, J. Tinkler, I. Viney, P. Wouters, J. Hill, and B. Johnson (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. London: HECFE.
- Zhao, T., Casu, B. Ferrari, A. (2010). 'The impact of regulatory reforms on cost structure, ownership and competition in Indian banking', *Journal of Banking and Finance*, **34**:246-54.

¹ Taylor also found that membership of the ‘1994 Group’ of ‘research-intensive’ universities had a strong impact (0.1% significance) on GPA scores in RAE2008.

² Following this reasoning, Sussex, Liverpool and Stirling (all with GPA= 3.125) would have replaced Birmingham, Loughborough and Warwick in Quartile 1 (see Table 1) if they had an assessor on the UoA19 panel.

³ The BBC, moreover, notes (following a review of web-sites) that Southampton still claims to be “In the top 1% of universities worldwide”, Liverpool states it is in the “while Queen’s University Belfast also portrays itself as “In the top one per cent of world universities.” (BBC News ‘Universities challenged on top 1% advert’, 8 June 2017).

⁴ According to REF2012 (p.77), ‘sustainability’ was the extent of future continuing capability ‘to support and develop research activity as defined in the quality levels, both within the submitted unit and the discipline more generally’; and ‘vitality’ was ‘the existence of a thriving, dynamic, and fully participatory research culture based on a clearly articulated research strategy, displayed both within the submitting unit and in its wider contributions, and in terms appropriate to the scale and diversity of the research activity that it supports’.

⁵ HEIs had three principal choices in REF2014: to maximise their ranking in performance league tables (which demanded submission of only those staff whose outputs were likely to be ranked as ‘world leading’ [4*]), to maximise their income (HEFCE applied a formula which took account of the quality of outputs and the number of staff returned in determining the funding awarded), or to maximise the number of staff submitted while promoting inclusivity. The latter would maximise neither income nor ranking.

⁶ Choice of which university submissions to subject to close reading was constrained by desire to avoid bias or conflict of interest between any author and a chosen institution (e.g., arising from prior employment, co-authorships or other collaborations).

⁷ This ‘superiority’ set contains some terms not included by *Microsoft Word*. Q1 institutions had more achievements to boast about and were better placed to provide evidence in support of such clichés.

⁸ A reviewer has drawn attention to the possibility that quality of research environment is endogenously determined. We do not infer causality. Nor do we offer any opinion about whether Q1 universities are highly ranked because they used language with particular style characteristics or whether they use language with such characteristics because they are highly ranked.

⁹ In contrast, the other four HEIs which were top ranked in ‘Environment’ (GPA = 4) submitted more than 75% of their eligible staff (http://charteredabs.org/wp-content/uploads/2015/03/refgpa_x_intensity.pdf).

¹⁰ Calculations of payments to individual universities were prepared from publicly available information by three staff members of one UK HEI, acting independently: an accountant, a Director of Finance, and an Associate Dean Research. All are highly conversant with REF-based funding allocation protocols.

¹¹ There is no precise mapping between UoAs and underlying school-departmental structures. This is evident in Economics, where many Economics Departments chose to submit to the B&M panel rather than the Economics

and Econometrics panel. For this reason, Pidd and Broadbent (2016, p.579) caution that care should be taken in interpreting staffing numbers.

¹² Pidd and Broadbent (2016, p.576) acknowledge that while some HEIs only submitted a small portion of their staff, they nonetheless submitted all their PGRs – even those PGRs whose supervisors were not entered, thus overinflating the declared number of PGRs per FTE.

¹³ The Herfindahl-Hirschman Index has been used to measure concentration in a variety of instances, including; household wealth and income concentration, market concentration in industrial and banking markets (Rhoades, 1993; Zhao *et al.*, 2010), competitive balance in professional team sports (Owen *et al.*, 2007), and citation concentration (Lariviere, 2009). We contend the index could be employed to measure the concentration of outputs under the REF scenario envisaged by Stern (2016) and HEFCE (2016).

¹⁴ The close reading results reported for ‘Environment’ are similar to those for UoA20, Law. The Chair of the Law sub-panel reported that strong submissions were characterised by clarity of structures and processes, clear descriptions of strategic links with context, and strong evidence to support claims (Douglas, 2015). Weaker submissions presented ‘activity rather than strategy’ and included inadequate evidence (e.g., focussing on staff recruitment rather than research completions). One distinctively different feature was that better submissions in Law provided ‘imaginative and vibrant research support with detailed thought given to supporting staff at all stages of their careers’ (Douglas, 2015, slide 22).